Optimal wave focusing in acoustic media

Farhad Bazargani[1], Roel Snieder[1] & Jon Sheiman[2]
[1] *Center for Wave Phenomena, Colorado School of Mines, Golden, CO 80401, USA*
[2] *Shell International Exploration and Production*

ABSTRACT
Focusing waves inside a medium has applications in various science and engineering fields, e.g., in medicine, nondestructive evaluation, ocean acoustics, and geophysics. The goal in focusing is to concentrate the wave energy at a specific time and location inside a medium. Various techniques have been devised and used to achieve this goal. Time-reversal is a method that is routinely used to focus acoustic and seismic waves. One important geophysical application of time-reversal focusing is in seismic source imaging. However, the method is not optimal for source imaging when acquisition is incomplete. Here, we propose a new technique wherein wave focusing is cast as an optimization problem. The new technique mitigates the adverse effects that incomplete acquisition has on time-reversal focusing and source imaging.

Key words: source imaging, focusing, optimization, time-reversal, acoustic emission

1 INTRODUCTION

The objective in wave focusing is to determine the waveforms that, when transmitted through a medium, create a wavefield that concentrates at a specific time and location. Wave focusing is conceptually related to the problem of imaging, and hence finds important applications in areas such as seismology and exploration geophysics.

Several methods for focusing have been devised, including those based on inverse scattering (Haddadin and Ebbini, 1998; Broggini et al., 2012; Behura et al., 2012), phase conjugation (Parvulescu, 1961), and time-reversal (Fink, 1997). Time-reversal (TR) is a well-established focusing technique that is robust and effective in heterogeneous media. The method relies on the time-reversal invariance of the wave operator and spatial reciprocity (Fink et al., 2002; Snieder, 2004).

A time-reversal mirror (TRM) is an array of transducers, each capable of recording, time-reversing (last-in first-out), and retransmitting signals into the medium. The TR process consists of two basic steps (Figure 1). In the first step, the wavefield generated by a source in the medium is recorded using a TRM. (To completely reconstruct the wavefield by the TR method, one needs a closed TRM surrounding the source.) In the second step, the recorded waveforms are time-reversed, retransmitted through the medium, and propagated back to refocus at the original source location. In a dissipative medium, time-reversal invariance is not satisfied. Spatial reciprocity alone, nevertheless, explains the robustness and efficiency of the TR process in many applications involving dissipative media (Fink, 2006).

TR focusing has been implemented in a variety of physical scenarios. It is applicable to both physical and numerical (back-propagation) experiments. In both, one deals with propagation of a time-reversed field, but the propagation is real in a physical problem and simulated in a back-propagation numerical problem (Fink, 2006). Applications involving physical TR include medical imaging (Robert and Fink, 2008), lithotripsy, underwater acoustics, and nondestructive testing (Fink, 1997; Larmat et al., 2010).

TR back-propagation methods are applied in key areas of geophysics on both global and exploration scales. In global seismology, TR techniques are used for studying earthquake source mechanism and location, for monitoring nuclear explosions, and in environmental applications of geophysics (Lu, 2002; Larmat et al., 2006, 2010). In exploration seismology, TR focusing is used in microseismic event location (McMechan, 1982; Lu and Willis, 2008; Xuan and Sava, 2010) and reservoir monitoring (Shapiro, 2008), in salt-flank imaging and redatuming seismic data (Lu, 2002), and in reversed time migration (McMechan, 1983; Berkhout, 1997; Schuster, 2002).

Despite these broad applications, the TR process has important theoretical and practical limitations. In theory, for a broadband pulse emitted by an ideal point

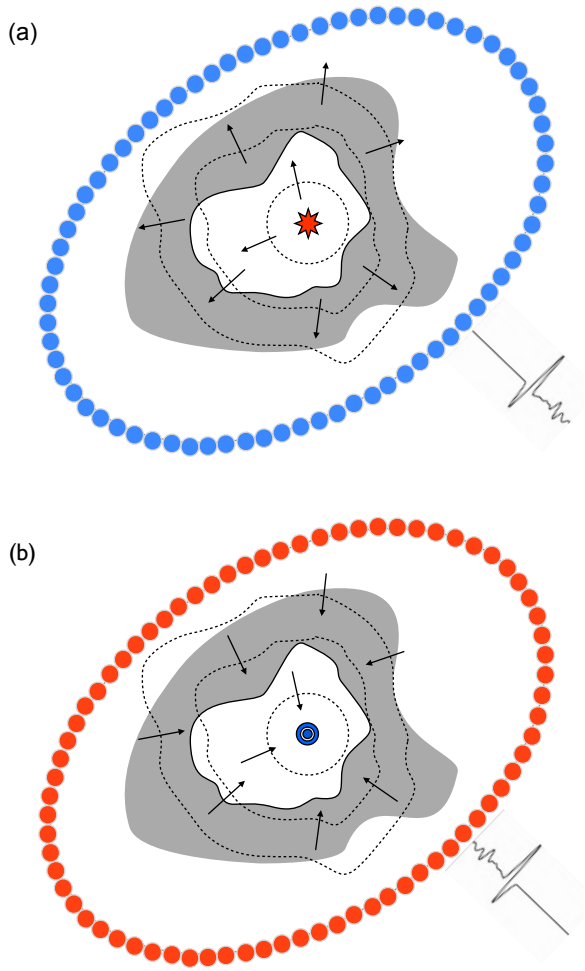

Figure 1. Illustration of a time-reversal experiment. (a) Forward propagation step: waves excited by a source travel through the complex medium and are recorded at stations marked as circles. (b) Back-propagation step: the recorded signals are reversed in time and rebroadcast into the medium at the corresponding stations. The waves then propagate through the medium and converge on the original source location.

source, the returning field refocuses on a spot with dimensions on the order of the smallest wavelength (Abbe diffraction limit). This is because evanescent waves containing source details smaller than the involved wavelengths cannot be sensed in the far-field. The loss of this information causes the resolution of the process to be bounded by the diffraction limit (Fink, 1997).

In practice, the wavefield is sampled at spatially sparse and limited locations. Also, it is not practical to surround the source with a full-aperture TRM, so a finite-aperture TRM is used instead. This incomplete acquisition results in a distortion in the shape of the point-spread function (Fink, 2006).

Another problem is that in real applications of TR-focusing, the media are dissipative and time-reversal invariance of the wave equation does not hold valid in dissipative media. However, as shown by Fink (2006), even in a dissipative medium, the TR process always maximizes the output amplitude at the focal time although it does not impose any constraints on the field around the focus. For example, side lobes can be observed around the source.

Several studies have been devoted to investigate these limitations and alleviate their effect to improve the TR procedure. Tanter et al. (2000, 2001), Aubry et al. (2001), and Montaldo et al. (2003) present the spatio-temporal inverse-filter method, a new focusing technique based on the inverse of the wave propagator between a source and elements in a TRM. For a lossless medium, the inverse filter method yields the same result as that of the TR method, but in the presence of attenuation, the spatio-temporal inverse filter methods are more effective than is the TR method.

Research on the connection between medium complexity and the size of the focal spot has shown a direct relationship between the complexity of the medium and the resolution in TR focusing; the more complicated the medium between the source and the TRM, the sharper the focus (Blomgren et al., 2002; Fink, 2008; Vellekoop et al., 2010). This is because a finite-aperture TRM acts as an antenna that uses complex environments to appear wider than it actually is, resulting in a focusing capability that is less dependent on the TRM aperture. In media made of random distribution of sub-wavelength scatterers, a time-reversed wave field can interact with the random medium to regenerate not only the propagating but also evanescent waves required to refocus below the diffraction limit (super-resolution). Schuster et al. (2012) demonstrate a method that uses evanescent waves generated by scatterers in the near-field region of seismic sources to achieve super-resolution.

We propose an alternative approach to wave focusing wherein the problem is cast as an optimization problem. As discussed above, the mainstream focusing methods currently in use are not optimal for real problems because of imperfect acquisition, attenuation, and the diffraction limit. The motivation for this research is to improve upon the existing techniques especially where such techniques do not perform optimally in dealing with the limitations posed on wave focusing by incomplete acquisition.

The organization of the paper is as follows: We first lay out the theoretical foundation of our new approach to wave focusing (section 2) and show how to design waveforms that will optimally focus at a desired known location and time inside a medium. Next we discuss how the new method, named Backus-Gilbert focusing (BG) is connected to some other focusing techniques like time-reversal and show that such other techniques are special cases of the more general solution to wave focusing that our approach provides (section 3). In the

next step, we show how, with a modicum of modification and reinterpretation, the BG theory can be adapted for application in source imaging problems (section 4). We start by adapting the BG theory for imaging a single point source, and then in appendix A we generalize the theory for an arbitrary spatio-temporally distributed source without knowing the properties of the source itself. Section 5 is devoted to two simple but representative numerical examples that demonstrate the application of the BG method in imaging of a point source and then a dipole source. In section 6 we discuss various aspects of the method and elaborate further on some explicit and implicit assumptions that are used in the construction of the new method and discuss their significance. Finally, in appendix B, we show an alternative time-domain formulation of the idea that underlies the BG method.

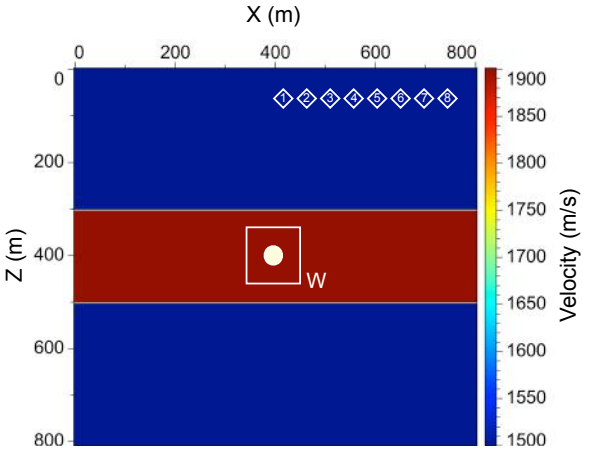

Figure 2. Acoustic velocity model and the configuration of the numerical experiments of section 5. The white diamonds show the location of the stations, the white square W depicts a small subset of the medium that contains the focusing target represented by the white circle.

2 FOCUSING AS AN OPTIMIZATION PROBLEM

Consider an acoustic medium with known velocity and N receiver stations at distinct locations $\mathbf{x}_i$ wherein each station records signals for T seconds for times $0 < t < T$. Suppose that at some time $\tau \in [0, T]$ an impulsive point source at location $\mathbf{x} = \boldsymbol{\xi}$ goes off and generates an acoustic wavefield that is eventually sampled by the N receivers as $d_i(t)$ for $i = 1, 2, ..., N$.

In the time-reversal method, we focus acoustic energy at $\mathbf{x} = \boldsymbol{\xi}$ and at $t = T - \tau$ by rebroadcasting the shifted time-reversed signals $d_i(T - t)$ at each station. However, because of the incomplete acquisition geometry of the receivers, the focus created by this TR process is suboptimal.

Our objective is to design signals such that, upon transmission, focus optimally at the target location $\boldsymbol{\xi}$ and at the time $T - \tau$. To achieve the best spatio-temporal focus, each station must work in concert with the others by injecting a signal that is tailored in amplitude and shape according to the medium properties, location of the focusing target, and the geometry of the stations.

If we denote the signal injected by the station at $\mathbf{x}_i$ as $a_i(t)$, then the superposed acoustic scalar wavefield recorded at an arbitrary location $\mathbf{x}$ inside the medium is

$$\phi(\mathbf{x}, t) = a_i(t) * G(\mathbf{x}, t; \mathbf{x}_i, 0), \quad \forall\, i \in \{1, 2, ...N\}, \qquad (1)$$

where we have used the Einstein's notational convention for summation over repeated indices, $*$ denotes the convolution operator, and $G(\mathbf{x}, t; \mathbf{x}_i, 0)$ is the Green function (impulse response) recorded at location $\mathbf{x}$ and corresponding to an impulsive source at time $t = 0$ and location $\mathbf{x}_i$. Note that if the medium is known, these Green functions can be computed.

Convolution in the time domain corresponds to multiplication in the frequency domain. Therefore, considering the problem in the frequency domain, using the Fourier convention $f(\mathbf{x}, \omega) = \int f(\mathbf{x}, t)\, \exp(i\omega t)\, dt$, each frequency component of the wavefield $\phi(\mathbf{x}, t)$ in equation 1 can be restated as a weighted sum of the corresponding frequency component $G(\mathbf{x}; \mathbf{x}_i, \omega)$ of the Green functions, i.e.,

$$\phi(\mathbf{x}, \omega) = a_i(\omega)\, G(\mathbf{x}; \mathbf{x}_i, \omega), \quad \forall\, i \in \{1, 2, ...N\}, \qquad (2)$$

where the weights $a_i(\omega)$ are the Fourier components of the signal that must be injected by the station at $\mathbf{x}_i$.

The problem can now be restated as how to optimally determine the unknown weights $a_i(\omega)$ in equation 2 so that the superposed field ϕ in the time domain focuses at a desired location of focus $\boldsymbol{\xi}$ and at a desired time of focus $T - \tau$. Put another way, the goal is to have $\phi(\mathbf{x}, t)$ as close as possible to $\delta(\mathbf{x} - \boldsymbol{\xi})\delta(t - T + \tau)$, where δ denotes the Dirac delta function. In the frequency domain, this goal can be achieved if we let each frequency component $\phi(\mathbf{x}, \omega)$ approach $e^{i\omega(T-\tau)}\delta(\mathbf{x} - \boldsymbol{\xi})$ by minimizing an objective function defined as

$$J = \int_W \left|\phi(\mathbf{x}, \omega) - e^{i\omega(T-\tau)}\delta(\mathbf{x} - \boldsymbol{\xi})\right|^2 d\mathbf{x}, \qquad (3)$$

where W denotes a subset of the medium that contains the target location (Figure 2).

Inserting equation 2 in objective function 3 and minimizing with respect to $a_i(\omega)$ results in a linear system of equations of the form

$$\boldsymbol{\Gamma}(\omega)\mathbf{a}(\omega) = e^{i\omega(T-\tau)}\mathbf{g}^*(\omega), \qquad (4)$$

where $\boldsymbol{\Gamma}$ is an $N \times N$ Gram matrix (Parker, 1994) defined

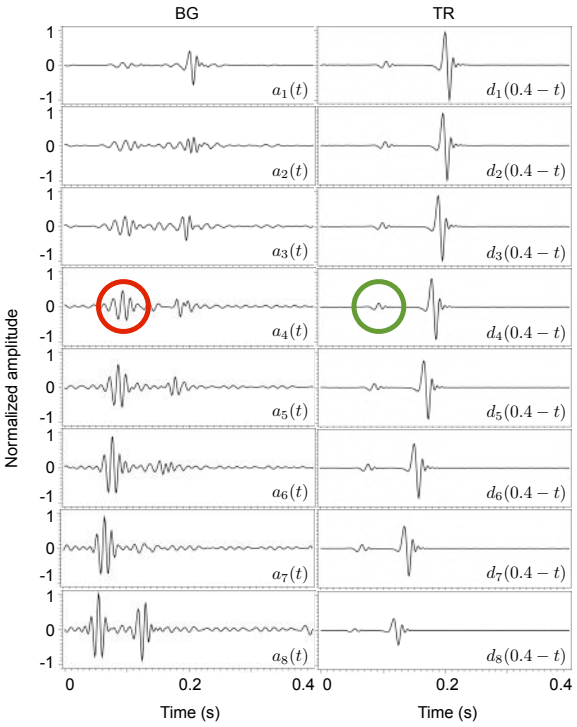

Figure 3. Optimized signals computed using the BG method (left column) associated with receivers 1 to 8, and the corresponding time-reversed data (right column). The weak reflected energy (green circle) in the TR trace is amplified (red circle) in the optimized BG trace.

by its elements as

$$\mathbf{\Gamma}_{ij}(\omega) = \int_W G(\mathbf{x};\mathbf{x}_i,\omega)\, G^*(\mathbf{x};\mathbf{x}_j,\omega)\, d\mathbf{x}, \qquad \forall\, i,j \in \{1,2,...,N\}, \tag{5}$$

$\mathbf{a}(\omega)$ is an $N \times 1$ vector that is the unknown of the equation, $\mathbf{g}$ is an $N \times 1$ vector with elements

$$g_i(\omega) = G(\boldsymbol{\xi};\mathbf{x}_i,\omega), \quad \forall\, i \in \{1,2,...,N\}, \tag{6}$$

and the superscript $*$ denotes complex conjugation. The linear system in 4 can be solved for the vector $\mathbf{a}(\omega)$ for each frequency separately. These $\mathbf{a}(\omega)$ vectors constitute the Fourier coefficients for the signals that the stations must transmit to obtain an optimal focus at location $\boldsymbol{\xi}$ and time $T-\tau$.

The condition that is implied by minimizing the objective function 3 is known as the *deltaness criterion* in the context of the method of Backus and Gilbert (BG) in inverse theory (Backus and Gilbert, 1968). A useful description of this method is provided by Aki and Richards (1980) and Aster et al. (2012). Hence, we adopt the term Backus-Gilbert focusing for the method described above for designing optimal signals for wave focusing.

3 CONNECTION WITH TIME-REVERSAL AND DECONVOLUTION

The Backus and Gilbert focusing method (BG), introduced in section 2, provides a more general solution to the wave-focusing problem compared to the other focusing techniques such as TR. In fact, there is a close mathematical relationship between these methods. To see the connection between BG and TR, let us replace $\mathbf{\Gamma}$ in equation 4 by the identity matrix I to get

$$\mathbf{a}(\omega) = e^{i\omega(T-\tau)}\mathbf{g}^*(\omega). \tag{7}$$

The complex conjugation and multiplication of $\mathbf{g}$ in equation 7 by $e^{i\omega(T-\tau)}$ for all frequencies amounts to in-place time-reversal of the corresponding signals $G(\boldsymbol{\xi}, t-\tau;\mathbf{x}_i,0)$ in the time domain. In other words, using the identity matrix as the crudest approximation for $\mathbf{\Gamma}$ results in the new system of equations 7 which describes exactly the same process as time-reversal in the time domain. Replacing $\mathbf{\Gamma}$ by the identity matrix amounts to ignoring the cross-talk between stations and having each station work independently to inject the time-reversed Green functions. Therefore, we might say that TR is a special cases of the more-general BG with a gross approximation of $\mathbf{\Gamma}$ as identity matrix.

In a similar way, we can show that BG is related to the deconvolution method (DC) presented by Ulrich et al. (2012). To show this relationship, let us set the off-diagonal elements of $\mathbf{\Gamma}$ equal to zero ($\mathbf{\Gamma}_{i,j} = 0$ for $i \neq j$). Solving the system of equations 4 for $\mathbf{a}$ gives

$$a_i(\omega) = \frac{e^{i\omega(T-\tau)}G^*(\boldsymbol{\xi};\mathbf{x}_i,\omega)}{\int_W G(\mathbf{x};\mathbf{x}_i,\omega)\, G^*(\mathbf{x};\mathbf{x}_i,\omega)\, d\mathbf{x}}, \qquad \forall\, i \in \{1,2,...,N\}. \tag{8}$$

In DC, the same frequency components of the signals to be back-propagated for focusing are computed as

$$a_i(\omega) = \frac{e^{i\omega(T-\tau)}G^*(\boldsymbol{\xi};\mathbf{x}_i,\omega)}{G(\boldsymbol{\xi};\mathbf{x}_i,\omega)\, G^*(\boldsymbol{\xi};\mathbf{x}_i,\omega) + \epsilon}, \qquad \forall\, i \in \{1,2,...,N\}, \tag{9}$$

where ϵ is a regularization term, on the order of $|G(\boldsymbol{\xi};\mathbf{x}_i,\omega)|^2$, added for stability of the solution. Notice the similarity between equations 8 and 9; the numerators on the right hand side of both equations are the same, and the denominators are similar except for the integration over the spatial element in equation 8 and the regularization term ϵ in equation 9. Therefore, keeping only the diagonal elements of $\mathbf{\Gamma}$ reduces BG to a method very similar to DC.

It is important to note that all elements of the Gram matrix $\mathbf{\Gamma}$, not just the diagonal elements, hold crucial information about the configuration of the wave-focusing experiment, i.e., the relative positions of the stations with respect to the propagation medium and the focusing target. Each element of $\mathbf{\Gamma}$ plays a role in determining how the stations must work together to inject the signals that achieve the optimum focusing at

the target. The function of the off-diagonal elements of $\mathbf{\Gamma}$ is to adjust the signal emitted by each station with respect to the signal of the other stations for the optimum focusing. Whereas in TR these off-diagonal elements are ignored by the crude approximation $\mathbf{\Gamma} = \mathbf{I}$ and therefore using the full $\mathbf{\Gamma}$ in BG can improve the TR focusing result.

4 APPLICATION IN SOURCE IMAGING

Since the beginning of modern seismology, understanding earthquake sources has been a focus of research. More recently, exploration geophysicists have become interested in studying the source of the microearthquakes that are generated during hydraulic fracturing of rocks (Rentsch et al., 2007). Techniques based on time-reversal focusing are now routinely used for seismic event location and source imaging (Larmat et al., 2010). The effectiveness of such techniques is, therefore, constrained by the same limitations (e.g., imperfect acquisition) that bound the efficacy of TR focusing. This means that the BG methodology proposed here to enhance TR focusing can be useful in source-imaging applications.

In a source-imaging problem, we would like to focus the energy of the wavefield, that is sampled (often sparsely and incompletely) as data by a limited number of receivers, back to its origin. When source imaging is viewed as a focusing problem, the location of the source is considered as the focusing target, and the activation time of the source is considered as the time of focus.

The formulation of BG presented in section 2 seems to require exact knowledge of the target location $\boldsymbol{\xi}$ and time τ. This means that in utilizing BG for source imaging, where the location of the source (focusing target $\boldsymbol{\xi}$) and its activation time (focusing time τ) are not known a priori, equation 4 may not be used directly. Nevertheless, as we show below, with a slight modification and reinterpretation of the BG formulation, explicit knowledge of the source location and time can be rendered unnecessary. In short, the BG method can be used in source imaging, because the required information about the source is encoded and implicitly available in the recorded data.

Let us begin by assuming that the source, that we intend to image, is an impulse $\delta(\mathbf{x} - \boldsymbol{\xi})\,\delta(t - \tau)$. In this context, the data $d_i(t)$ recorded by a receiver at $\mathbf{x}_i$ due to our impulsive source is the Green function $G(\mathbf{x}_i, t; \boldsymbol{\xi}, \tau)$. The reciprocity of the Green function allows for expressing this data in the frequency domain as

$$d_i(\omega) = e^{i\omega\tau}\, G(\boldsymbol{\xi}; \mathbf{x}_i, \omega). \tag{10}$$

Now, using 6 and 10, equation 4 can be rewritten as

$$\mathbf{\Gamma}(\omega)\, \mathbf{a}(\omega) = e^{i\omega T} \mathbf{d}^*(\omega), \tag{11}$$

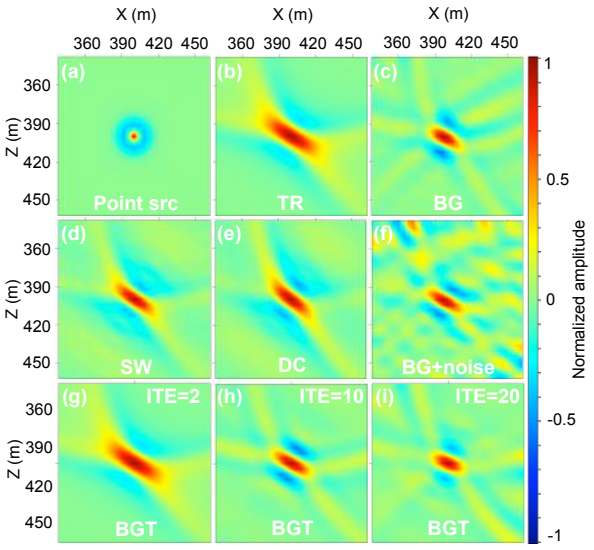

Figure 4. Exact point source (a) and its images produced by time-reversal (b), Backus-Gilbert (c), time-reversal-plus-spectral-whitening (d), deconvolution (e), Backus-Gilbert using data contaminated by noise with $S/N = 1$ (f), iterative Backus-Gilbert in the time domain (BGT) after 2 iterations (g), BGT after 10 iterations (h), and BGT after 20 iterations (i)

where $\mathbf{d}(\omega)$ is an $N \times 1$ vector with elements defined by equation 10.

Note that the right hand side of this equation is now completely known. The significance of equation 11 is that it allows us to use the BG formalism for imaging an impulsive source with no knowledge of the source time and location. The actual location and time of the source can be found, eventually, by solving equation 11 for the optimized signals $\mathbf{a}(\omega)$, injecting them by the receivers, and scanning the resulting wavefield for the source image. In other words, after solving equation 11 for the optimized signals, the procedure for finding the source time and location (imaging the source) would be exactly similar to the usual practice in time-reversal source imaging.

The argument above relied on our initial assumption that the source was impulsive. However, as shown in appendix A, this argument can be generalized to hold true for any arbitrary spatio-temporal acoustic source, meaning that equation 11 can be used regardless of the source being impulsive or not.

5 NUMERICAL EXPERIMENTS

Here we test the ideas presented above by performing two numerical source-imaging experiments. We first apply the BG method to image a point source and compare the result with the same image produced by other tech-

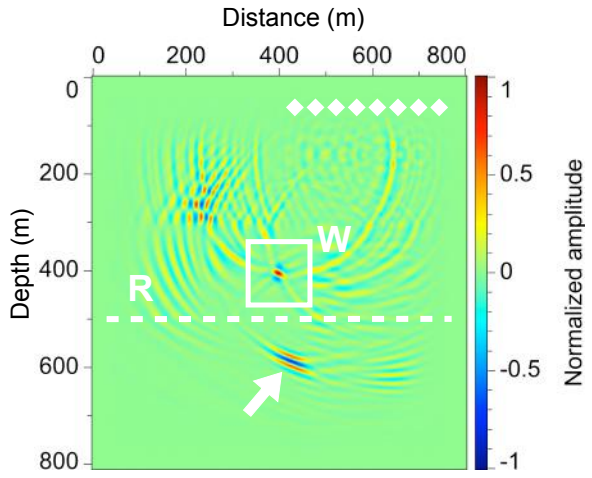

Figure 5. BG wavefield at the time of focus. The area within W is the same as Figure 4c. The dashed line R denotes the location of the reflector and the diamonds represent the 8 receivers. The white arrow points to the energy boosted by BG such that its reflection from R illuminates the target from below.

niques such as TR, DC, and time-reversal plus spectral-whitening (SW). Spectral whitening is a simple method for boosting the weak frequencies of a signal by dividing the components of the frequency band by their magnitudes such that the resulting spectrum is flattened (whitened). We next use BG to image a dipole, as an example of a spatially distributed source, and see how the resulting image compares with the same image produced by TR.

The configuration of the experiments is shown in Figure 2. The eight diamonds represent the receivers, the white circle depicts the source location, and the white square shows the optimization window W used in the definition of the objective function 3 in the formulation of the BG scheme. The velocity model used for wave propagation is a heterogeneous 2D model (Figure 2) consisting of three layers. Wave propagation is simulated using an explicit finite-difference approximation of the 2D acoustic isotropic wave equation with absorbing boundary condition.

In the first experiment, to simulate the data $d_i(t)$ we generate a source wavefield by injecting a Ricker wavelet with peak frequency of $64\,Hz$ at time $\tau = 50\,ms$ and location $\boldsymbol{\xi} = (400\,m, 400\,m)$ within W. We sample this source wavefield by the receiver at $\mathbf{x}_i$ for $T = 0.4\,s$. After simulating the data, for the rest of the experiment, we pretend that we do not know the actual location and time of the source.

To form the Gram matrix $\mathbf{\Gamma}$ in equation 11, we require the Green functions $G(\mathbf{x}, t; \mathbf{x}_i, 0)$, $\forall\, i \in \{1, 2, ..., N\}$. We approximate these Green functions by injecting a band limited spike with frequencies between $5\,Hz$ and $150\,Hz$ at each receiver location $\mathbf{x}_i$ and propagate the wavefield for $T = 0.4\,s$. These wavefields are then Fourier transformed to the frequency domain and used in equation 5 to compute the elements of the 8×8 complex matrix $\mathbf{\Gamma}$ independently for all frequencies within the bandwidth of the experiment. In computing the integral of equation 5, the oscillatory integrand is tapered near the edges of integration window W to avoid dominant contribution from the end points.

At this point, we can form the system of equations 11 for each frequency independently and solve the system for $a_i(\omega)$, the Fourier coefficients of the optimized signals $a_i(t)$. These optimized signals are then broadcast by the receivers to generate the optimal wavefield $\phi(\mathbf{x}, t)$ that will focus to create the image of the source at location $\boldsymbol{\xi} \in W$ and at time $T - \tau$. As with TR, the last step is to scan the wavefield $\phi(\mathbf{x}, t)$ to detect and extract the source image. (The source image can be detected using, for example, its high energy.) After detecting the source image, the actual values of $\boldsymbol{\xi}$, τ, and also the spatio-temporal characteristics of the source can be inferred from that image.

Figures 3 and 4 summarize the results of the first experiment. Figure 3 shows the normalized optimal signals $a_i(t)$ and the normalized time-reversed data $d_i(t)$ for receivers 1 through 8. The signals in each columns have been normalized by dividing the amplitude of each sample by the maximum absolute value of the amplitude of all traces in the same column. The optimization process has produced signals (left column) that are different from their corresponding time-reversed data (right column) in both amplitude and shape. For example, the small amplitude events in the time-reversed traces (e.g., the energy encircled in green) correspond to the reflection energy that is reflected from the discontinuity at $500\,m$ in Figure 2. Note how the same reflected events (e.g. the energy encircled in red) are amplified by BG in the optimally computed signals.

Figure 4a shows the exact source in the first experiment. More specifically, it depicts the portion of the source wavefield enclosed within W at the activation time of the point Ricker wavelet that was used to simulate the source. Figure 4b shows the image of this source produced by TR. Instead of a compact and round spot, the TR image of the source is distorted. This distortion is a consequence of the incomplete acquisition geometry of the receivers. For recovering a spot-like image of the point source, TR requires a balanced illumination of the target from all angles. However, in this TR experiment, target illumination is imbalanced and limited to the small angle subtended by the first and last receivers.

The source image produced by BG is shown in Figure 4c. BG has outperformed TR in achieving a more compact and spot-like image of the point source. This improvement is mostly the result of augmentation of the poor target illumination by BG. Figure 5 shows the entire BG wavefield $\phi(\mathbf{x}, t)$ at the time of focus. (The

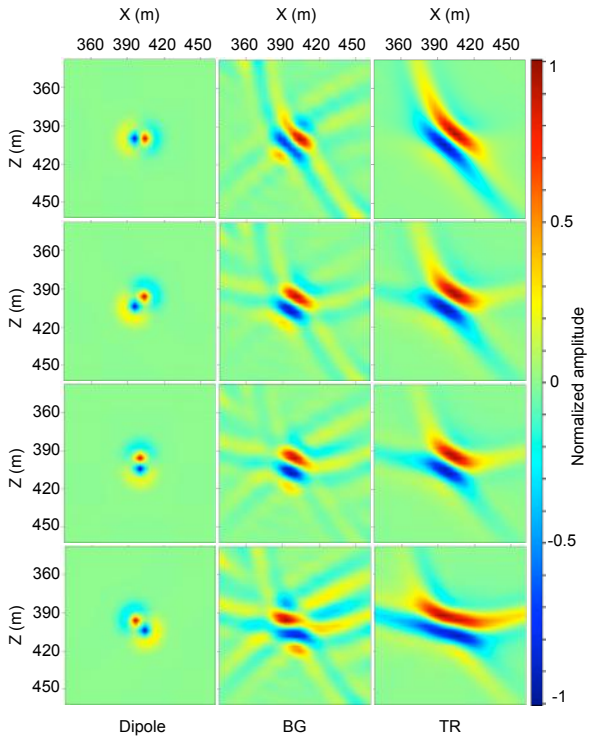

Figure 6. Comparison of BG ans TR in imaging a dipole as an example of a distributed source with anisotropic radiation pattern. Snapshots of the wavefield associated with dipole sources with different orientations (left column) and there corresponding BG images (middle column) and TR images (right column).

portion enclosed by W here is the same as Figure 4c). The strong burst of energy denoted by the white arrow in Figure 5 is the result of propagating the amplified reflected events in the optimized signals shown in Figure 3. When the optimized signals are propagated, this strong burst of energy travels in advance and part of it, after bouncing off the reflector at $Z = 500\,m$, illuminates the target from below. Of course, BG did not create this energy out of nowhere. The energy is also present in the TR experiment, but it is much weaker. All that was done by BG was to detect this weak energy and amplify it in order to balance the target illumination. Effectively, this is equivalent of using the reflector at $Z = 500\,m$ as an acoustic mirror in order to boost the illumination angles.

For the sake of generality, we also imaged the point source using using SW (Figure 4d) and DC (Figure 4e) . Although the results by SW and DC are a slight improvement over TR (Figure 4b) they are both significantly less localized in space than the BG image (Figure 4c).

In the second experiment, we keep the configuration of the source and receivers and velocity the same as in the first experiment. However, instead of a point source for simulating the data, we use a dipole with time dependence given by a Ricker wavelet. The remaining steps follow exactly the same as before. Unlike a point source that has an isotropic radiation pattern, a dipole's radiation pattern is directional. Consequently, in a source imaging experiment with incomplete acquisition, the arrangement of the receivers relative to the nodal planes and lobes of the source radiation pattern becomes an important factor. For this reason, we repeat the second experiment for 4 different orientations of the source dipole, i.e., for $\theta = 0$, $\pi/4$, $\pi/2$, and $3\pi/4$, where θ denotes the angle of the dipole orientation measured with respect to the horizontal.

Figure 6 summarizes the outcome of these experiments. The left column shows snapshots of the wavefield associated with dipole sources with different orientations at the time of source activation. The middle column shows the image of these dipole sources as obtained by BG, and the right column depicts the image of the same dipole sources produced by TR. The source images obtained by BG are more focused than their corresponding TR images. Besides, compared with TR and for all dipole orientations, BG shows better delineation of the orientation of the dipole source than TR.

6 DISCUSSION

Knowledge of the medium (e.g. velocity and density) is the most significant assumption made in formulating the BG method. This information is needed for computing the Green function for each receiver location. These Green functions are in turn used for computing the Γ matrix.

As in any imaging method, the accuracy of the velocity model is very important in BG. Using an inaccurate velocity model can generally cause the image of the source to be created at a wrong time or location. More specifically, in situations like the examples of section 5, the BG method relies on amplifying the weak reflection energy in data in order to mitigate the adverse effects of incomplete acquisition. In such situations, for BG to effectively balance the illumination, the velocity model must be accurate and contain the reflectors that are associated with the main reflection energy in the data. Otherwise, BG will not be as effective as it is portrayed by examples of section 5 and, depending on the fidelity of the velocity model in predicting the reflections and how far it is from the true velocity model, the BG image of the source will be degraded and approach that of TR. (Recall that TR corresponds to the crude approximation $\mathbf{\Gamma} \approx I$ and therefore is less sensitive to the inaccuracies in the velocity model.)

Apart from the velocity model, we also need an estimate of the source location so that we can be certain that the source is somewhere within a limited area W. In practice, such an estimate is usually available or

can be obtained through using common source location techniques such as TR. The size of W is an important factor in determining the effectiveness of BG as a source imaging method. This is because W is used directly in the definition of the objective function 3 upon which we based the formulation of the BG method. With a limited number of receivers, the smaller W (the smaller the uncertainty in the location of the source) is, the easier it is to minimize the BG objective function. However, there is a lower limit on how small W can be to obtain best focusing results. Based on our tests (not shown here) the optimum focusing is achieved when the size of W is 2 to 3 times the dominant wavelength in the data.

The presence of noise in recorded data is an important matter that needs careful attention in applying BG in source imaging. In our numerical examples of section 5, we did not include noise. Applying the BG method for source imaging hinges upon the validity of equation 10 which allows us to form and solve the system of equations 11 without explicit knowledge of the source location, time, and spatio-temporal character. In a real source imaging scenario, data is always contaminated with noise and, therefore, equation 11 must be modified as

$$\mathbf{\Gamma}\,(\omega)\mathbf{a}(\omega) \approx e^{i\omega T}(\mathbf{d}^*(\omega) + \boldsymbol{\eta}^*(\omega)), \qquad (12)$$

with $\boldsymbol{\eta}$ denoting the noise vector and $\mathbf{d}$ the noise free data described by equation 10.

The stability of the solution $\mathbf{a}$ to equation 12 depends on the condition number of the matrix $\mathbf{\Gamma}(\omega)$, which itself depends on the configuration of the receivers, the properties of the medium, and the frequency for which the $\mathbf{\Gamma}(\omega)$ is being computed. If $\mathbf{\Gamma}(\omega)$ is ill-conditioned, then some regularization technique, e.g, truncated SVD can be used to find a stable solution to equation 12.

To test the robustness of the BG method in the presence of noise, and also to demonstrate the effect of additive noise to data, we repeated the first numerical source imaging experiment in section 5 for data contaminated with band-limited Gaussian random noise with the same bandwidth as the data and signal-to-noise ratio of 1. The result of this test is shown in Figure 4f. For this geometry, $\mathbf{\Gamma}(\omega)$ is not well-conditioned for low frequencies. This is because for lower frequencies (long wavelengths) the Green functions computed for adjacent receivers become similar with the result of $\mathbf{\Gamma}(\omega)$ becoming more singular. Therefore, to find a well-behaved solution, we regularized the system 12 using truncated singular value decomposition (TSVD) for low frequencies.

The formulation of the BG focusing presented in this paper has so far been in the frequency domain. However, it is possible to formulate the same optimization idea in the time domain as well. The details of the time-domain formulation are shown in appendix B. Solving the BG optimization problem in the time domain involves an iterative solution, where each iteration includes two wave propagation steps (see equation B12). In this respect, time-domain Backus-Gilbert focusing (BGT) resembles least-squares migration (Nemeth et al., 1999). Considering that this iterative process must be repeated for imaging each source, implementing the BG method in the time domain could be expensive.

Solving the BG optimization in the time domain or in the frequency domain are fundamentally the same problems; in both, we minimize the same objective function. This is a consequence of the Parseval's theorem (compare objective functions B1 and A4). To illustrate this point, we repeated the first numerical source imaging experiment in section 5 for imaging a point source with BGT. Figures 4g, 4h, and 4i show the result of BGT after 2, 10, and 20 iterations of the conjugated-gradient method, respectively. In early iterations (Figure 4g) the BGT focus is comparable to the TR focus (Figure 4b). However, by increasing the number of conjugate-gradient iterations, the resolution of the source image produced by BGT improves and after 20 iterations, the BGT focus (Figure 4i) approaches that of frequency-domain BG (Figure 4c).

Minor differences between the time- and frequency-domain BG focusing results must be attributed to the differences in numerical implementations of the two methods. While in the time-domain, we solve the BG optimization problem using an iterative scheme, in the frequency-domain, the BG optimization problem is solved directly by inverting the associated linear systems for each frequency. For this reason, adding noise to data can have a different effect on the frequency-domain BG than on the time-domain BG and depending on the method used for regularizing the inverse problem and the noise level, the difference between the results obtained by time-domain and frequency-domain BG can increase.

In the frequency-domain formulation, computing the matrix $\mathbf{\Gamma}$ requires simulating the wave propagation several times to model the Green function for each receiver location. However, a nice feature of the frequency-domain implementation is that for a fixed configuration of the receivers, $\mathbf{\Gamma}$ and its inverse have to be computed only once. (Recall that $\mathbf{\Gamma}$ on the left hand side of equation 11 depends only on the receiver geometry and the properties of the medium and the right hand side depends only on the recorded data.) Therefore, once $\mathbf{\Gamma}^{-1}$ is computed and stored, we can reuse it to solve equation 11 whenever a new source occurs. This means that for receivers with a fixed geometry, applying the BG in the frequency domain to image multiple sources can be very fast, inexpensive, and computationally efficient. However, note that in a case where the number of receivers are too large, and the number of sources to be imaged are small, it might be preferable to solve the BG problem in the time domain.

7 CONCLUSION

The ability of time-reversal (TR) methods to focus waves inside heterogeneous media is bounded by limitations such as attenuation, the diffraction limit, and imperfect acquisition. To go beyond the limitations caused by imperfect acquisition, we formulate wave focusing as an optimization problem. This technique, called the Backus-Gilbert (BG) method, can be easily applied in source imaging to obtain optimal images of an acoustic source.

The only requirements are accurate knowledge of the medium, and an estimate of the source location. The method makes no particular assumptions about the spatio-temporal character of the source. All the information that the method needs about the source is encoded in the recorded data.

Our numerical tests show that the Backus-Gilbert (BG) approach is capable of achieving a significantly better resolved image of the source compared to that achieved by TR. The advantage of BG over TR is most pronounced in source imaging experiments with limited aperture coverage and severely incomplete acquisition geometry. In fact, the aim of the method is mitigating the adversed effects of imperfect acquisition, and we emphasise that BG can not beat the diffraction limit as it ultimately relies on propagating band-limited signals from a limited number of stations to achieve an optimal focus.

The BG method can find applications in geophysics in both global and exploration scales. In global seismology, techniques based on TR focusing are routinely used to image earthquake sources (Larmat10). In exploration scale, geophysicists have interest in imaging and understanding the source mechanism of the micro earthquakes generated in hydraulic fracturing experiments. Often, in both earthquake source imaging and microseismic monitoring, the number and distribution of receivers is not enough to sample the wavefield properly. For example, in global seismology, most seismometer stations are on the continents and in microseismic surveys, the receivers are installed in one or two monitoring wells. In examples like these, TR can not perform optimally and BG could be a good candidate to compensate for the incomplete acquisition and poor aperture coverage.

The earth is, however, an elastic medium and therefore, to be able to apply BG in geophysical source imaging applications, we need to develop an equivalent elastic version of the BG method presented in this paper. We deal with this problem specifically in report CWP-792.

8 ACKNOWLEDGMENTS

This research was supported by the sponsor companies of the Consortium Project on Seismic Inverse Methods for Complex Structures. The software used to produce the results in the paper was written in Java, Jython with the use of libraries in the Mines Java Toolkit freely available at `https://github.com/dhale/jtk`.

REFERENCES

Aki, K., and P. Richards, 1980, Quantitative seismology theory and methods, vol. 2: Freeman and Company.

Aster, R., B. Borchers, and C. Thurber, 2012, Parameter estimation and inverse problems 2nd edition: Academic Press.

Aubry, J. F., M. Tanter, J. Gerber, J. L. Thomas, and M. Fink, 2001, Optimal focusing by spatio-temporal inverse filter. ii. experiments. application to focusing through absorbing and reverberating media: The Journal of the Acoustical Society of America, **110**, 48–58.

Backus, G., and F. Gilbert, 1968, The resolving power of gross earth data: Geophysical Journal of the Royal Astronomical Society, **16**, 169–205.

Behura, J., K. Wapenaar, and R. Snieder, 2012, Newton-Marchenko-Rose Imaging: Image reconstruction based on inverse scattering theory: Technical Report CWP-720, Center for Wave Phenomena, Colorado School of Mines.

Berkhout, A., 1997, Pushing the limits of seismic imaging, part i: Prestack migration in terms of double dynamic focusing: Geophysics, **62**, 937–954.

Blomgren, P., G. Papanicolaou, and H. Zhao, 2002, Super-resolution in time-reversal acoustics: The Journal of the Acoustical Society of America, **111**, 230–248.

Broggini, F., R. Snieder, and K. Wapenaar, 2012, Focusing the wavefield inside an unknown 1d medium: Beyond seismic interferometry: Geophysics, **77**, A25–A28.

Fink, M., 1997, Time-reversed acoustics: Physics Today, **50**, no. 3, 30–34.

———, 2006, Time-reversal acoustics in complex environments: Geophysics, **71**, 151–164.

———, 2008, Time-reversal waves and super resolution: Journal of Physics:Conference Series, **124**, 1–29.

Fink, M., W. Kuperman, J. Montagner, and A. Tourin, 2002, Imaging of complex media with acoustic and seismic waves: Springer-Verlag.

Haddadin, O., and E. S. Ebbini, 1998, Ultrasonic focusing through inhomogeneous media by application of the inverse scattering problem: Journal of the Acoustical Society of America, **104**, 313–325.

Larmat, C., R. A. Guyer, and P. A. Johnson, 2010, Time-reversal methods in geophysics: Physics Today, **63**, no. 8, 31–35.

Larmat, C., J. Montagner, M. Fink, and Y. Capdeville, 2006, Time-reversal imaging of seismic sources and application to the great sumatra earthquake: Geophysical Research Letters, **33**, L19312.

Lu, R., 2002, Time reversed acoustics and applications to earthquake location and salt dome flank imaging: PhD Thesis, Massachusetts Institute of Technology.

Lu, R., and M. E. Willis, 2008, Locating microseismic events with time reversed acoustics: A synthetic case study: SEG Technical Program Expanded Abstracts, 1342–1346.

McMechan, G. A., 1982, Determination of source parameters by wavefield extrapolation: Geophysical Journal of the Royal Astronomical Society, **71**, 613–628.

———, 1983, Migration by extrapolation of time-dependent boundary values: Geophysical Prospecting, **31**, 413–420.

Montaldo, G., M. Tanter, and M. Fink, 2003, Real time inverse filter focusing through iterative time reversal: The Journal of the Acoustical Society of America, **115**, 768–775.

Nemeth, T., C. Wu, and J. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.

Parker, R., 1994, Geophysical inverse theory: Princeton University Press.

Parvulescu, A., 1961, Signal detection in a multipath medium by M.E.S.S. processing: The Journal of the Acoustical Society of America, **33**, 1674.

Rentsch, S., S. Buske, S. Luth, and A. Shapiro, 2007, Fast location of seismicity: A migration-type approach with application to hydraulic-fracturing data: Geophysics, **72**, S33–S40.

Robert, J., and M. Fink, 2008, Greens function estimation in speckle using the decomposition of the time reversal operator: Application to aberration correction in medical imaging: The Journal of the Acoustical Society of America, **123**, 866–877.

Schuster, J. T., 2002, Reverse-time migration = generalized diffraction stack migration: 72nd Annual International Meeting, SEG, Expanded Abstracts, 1280–1283.

Schuster, J. T., S. Hanafy, and Y. Huang, 2012, Theory and feasibility tests for a seismic scanning tunnelling macroscope: Geophysical Journal International, **190**, 1593–1606.

Shapiro, S. A., 2008, Microseismicity a tool for reservoir characterization: EAGE Publications bv.

Snieder, R., 2004, A guided tour of mathematical methods for the physical sciences, 2nd ed.: Cambridge Univ. Press.

Tanter, C., J. L. Thomas, and M. Fink, 2000, Time reversal and inverse filter: Journal of the Acoustical Society of America, **108**, 223–234.

Tanter, M., J. F. Aubry, J. Gerber, J. Thomas, and M. Fink, 2001, Optimal focusing by spatio-temporal inverse filter. i. basic principles: Journal of the Acoustical Society of America, **110**, 37–47.

Ulrich, T., B. Anderson, C. Le Bas, P.Y. Payan, J. Douma, and R. Snieder, 2012, Improving time reversal focusing through deconvolution: 20 questions: Proc. Meet. Acoustics, **16**, 045015.

Vellekoop, I., A. Lagendijk, and A. Mosk, 2010, Exploiting disorder for perfect focusing: Nature Photonics, **4**, 320–322.

Xuan, R., and P. Sava, 2010, Probabilistic microearthquake location for reservoir monitoring: Geophysics, **75**, MA9–MA26.

APPENDIX A: OPTIMIZED IMAGING OF AN ARBITRARY ACOUSTIC SOURCE

In section 4, we discussed application of our method for imaging an impulsive source at an unknown time and location. Here, we show that this method is equally applicable for imaging any source function (not just impulsive) with an arbitrary spatio-temporal characteristic.

Let $s(\mathbf{x},t)$ denote the spatio-temporal source function defined over all space and time such that it can be nonzero only for $\mathbf{x}\in W$ and $t\in[0,T]$. We can think of the source function as a succession of spatially impulsive sources that are applied with strength $s(\boldsymbol{\xi},t)$ at each location $\boldsymbol{\xi}$ and write

$$s(\mathbf{x},t)=\int s(\boldsymbol{\xi},t)\,\delta(\mathbf{x}-\boldsymbol{\xi})\,d\boldsymbol{\xi}\,. \tag{A1}$$

The data recorded by a station at $\mathbf{x}_i$ associated with this distribution of impulsive sources can then be written as

$$d_i(t)=\int s(\boldsymbol{\xi},t)*G(\mathbf{x}_i,t;\boldsymbol{\xi},0)\,d\boldsymbol{\xi}\,, \tag{A2}$$

where $G(\mathbf{x}_i,t;\boldsymbol{\xi},\tau)$ is the Green function with the source at $\mathbf{x}=\boldsymbol{\xi}$. In the frequency domain, equation A2 becomes

$$d_i(\omega)=\int s(\boldsymbol{\xi},\omega)\,G(\mathbf{x}_i;\boldsymbol{\xi},\omega)\,d\boldsymbol{\xi}\,. \tag{A3}$$

Having defined the source function and the data associated with it, we can now lay out the BG optimization problem.

We assume that the scalar source wavefield at the time of focus is proportional to the source function $s(\mathbf{x},t)$, with unity as the proportionality constant. With this assumption, we define our goal as finding signals $a_i(t)$ such that the difference between the wavefield $\phi(\mathbf{x},t)=a_i(t)*G(\mathbf{x},t;\mathbf{x}_i,0)$ and the time-reversed source fucntion $s(\mathbf{x},T-t)$ is minimum.

We achieve this in by minimizing an objective function defined as

$$J=\iint_W|\phi(\mathbf{x},t)-s(\mathbf{x},T-t)|^2d\mathbf{x}\,dt, \tag{A4}$$

which can be expressed in the frequency domain as

$$J(\omega) = \int_W \mid a_i(\omega)\, G_i(\mathbf{x};\mathbf{x}_i,\omega) - \int e^{i\omega T}\, s^*(\boldsymbol{\xi},\omega)\, \delta(\mathbf{x}-\boldsymbol{\xi})\, d\boldsymbol{\xi} \mid^2 d\mathbf{x}, \tag{A5}$$

for each frequency. Minimization A5 with respect to $a_i(\omega)$ gives

$$\begin{aligned} a_j(\omega) \int_W G(\mathbf{x};\mathbf{x}_i,\omega) G^*(\mathbf{x};\mathbf{x}_j,\omega) d\mathbf{x} \\ = e^{i\omega T} \int s^*(\boldsymbol{\xi},\omega)\, G^*(\boldsymbol{\xi};\mathbf{x}_i,\omega)\, d\boldsymbol{\xi} \\ = e^{i\omega T} \int s^*(\boldsymbol{\xi},\omega)\, G^*(\mathbf{x}_i;\boldsymbol{\xi},\omega)\, d\boldsymbol{\xi}\,, \end{aligned} \tag{A6}$$

where in the last step we have used the reciprocity principle for the acoustic Green function. Using equation A3, equation A6 can be written as

$$a_j(\omega) \int_W G(\mathbf{x};\mathbf{x}_i,\omega)\, G^*(\mathbf{x};\mathbf{x}_j,\omega)\, d\mathbf{x} = e^{i\omega T}\, d_i^*(\omega)\,, \tag{A7}$$

or more concisely

$$\boldsymbol{\Gamma}(\omega)\, \mathbf{a}(\omega) = e^{i\omega T} \mathbf{d}^*(\omega)\,, \tag{A8}$$

where $\boldsymbol{\Gamma}$ is an $N \times N$ matrix with the same definition as given by equation 5 in section 2 and $\mathbf{d}(\omega)$ is an $N \times 1$ vector with elements defined as A3, which is the frequency component of the data recorded by station at $\mathbf{x}_i$.

APPENDIX B: TIME-DOMAIN FORMULATION OF THE BG METHOD

In section 2 we formulated the BG focusing method in the frequency domain. Here, we derive the time-domain formulation of the same idea.

Let $s(\mathbf{x},t)$ denote the source function defined over $\mathbf{x} \in W$ and $t \in [0,T]$. Where W and T, as defined in section 2, represent the optimization window and the maximum recording time, respectively. We assume that the scalar source wavefield at the time of focus is proportional to the source function $s(\mathbf{x},t)$, with unity as the proportionality constant. With this assumption, the goal of the optimization is to find signals $a_i(t)$, the signals the must be injected by a station at location $\mathbf{x}_i$, such that the difference between the wavefield $\phi(\mathbf{x},t) = a_i(t) * G(\mathbf{x},t;\mathbf{x}_i,0)$ and the shifted time-reversed source function $s(\mathbf{x},T-t)$ is minimum. We achieve this goal by minimizing an objective function defined as

$$J = \iint_W |\phi(\mathbf{x},t) - s(\mathbf{x},T-t)|^2 d\mathbf{x}\, dt. \tag{B1}$$

Substituting

$$\phi(\mathbf{x},t) = \int dt' a_i(t') G(\mathbf{x},t-t';\mathbf{x}_i,0) \tag{B2}$$

into B1, and minimizing J with respect to $a_n(t_m)$ for some m and n by setting $\partial J/\partial a_n(t_m) = 0$ gives

$$\begin{aligned} &\int dt \int_W d\mathbf{x}\, [G(\mathbf{x},t-t_m;\mathbf{x}_n,0) \\ &\int dt' G(\mathbf{x},t-t';\mathbf{x}_i,0) a_i(t')] \\ &= \int dt \int_W d\mathbf{x}\, G(\mathbf{x},t-t_m;\mathbf{x}_n,0) s(\mathbf{x},T-t). \end{aligned} \tag{B3}$$

Replacing t with τ, i with j, n with i, and t_m with t, equation B3 becomes

$$\begin{aligned} &\int d\tau \int_W d\mathbf{x}\, [G(\mathbf{x},\tau-t;\mathbf{x}_i,0) \\ &\int dt' G(\mathbf{x},\tau-t';\mathbf{x}_j,0) a_j(t')] \\ &= \int d\tau \int_W d\mathbf{x}\, G(\mathbf{x},\tau-t;\mathbf{x}_i,0) s(\mathbf{x},T-\tau) \\ &= \int d\tau \int_W d\mathbf{x}\, G(\mathbf{x}_i,\tau-t;\mathbf{x},0) s(\mathbf{x},T-\tau)\,, \end{aligned} \tag{B4}$$

where in the last step, we used of the reciprocity of the Green's function, $G(\mathbf{x},t;\mathbf{x}_i,0) = G(\mathbf{x}_i,t;\mathbf{x},0)$. This equation can be further simplified if we define a new function

$$s'(\mathbf{x},t) = s(\mathbf{x},T-t), \tag{B5}$$

and substitute it in equation B4 to get

$$\begin{aligned} &\int d\tau \int_W d\mathbf{x}\, [G(\mathbf{x}_i,\tau-t;\mathbf{x},0) \\ &\int dt'\, G(\mathbf{x},\tau-t';\mathbf{x}_j,0)\, a_j(t')] \\ &= \int d\tau \int_W d\mathbf{x}\, G(\mathbf{x}_i,\tau-t;\mathbf{x},0)\, s'(\mathbf{x},t)\,. \end{aligned} \tag{B6}$$

Now, we define the propagation operator L such that

$$L\,\mathbf{f}(t) = \int G(\mathbf{x},t-\tau;\mathbf{x}_j,0)\, f_j(\tau)\, d\tau, \tag{B7}$$

and the adjoint propagator $L^\dagger$ such that

$$L^\dagger\, \theta(\mathbf{x},t) = \iint_W G(\mathbf{x}_i,\tau-t;\mathbf{x},0)\, \theta(\mathbf{x},\tau)\, d\mathbf{x}\, d\tau,\ \forall i \tag{B8}$$

where $\mathbf{f}(t)$ is an arbitrary vector function of time with components $f_i(t)$, and $\theta(\mathbf{x},t)$ is an arbitrary space-time function. The adjointness of L and $L^\dagger$ defined above can be verified by showing

$$< Lf(t)\,,\, \theta(\mathbf{x},t) > = < f(t)\,,\, L^\dagger \theta(\mathbf{x},t) >\,, \tag{B9}$$

where the 1D and 2D inner products on the left and right hand side of the identity B9 are defined as

$$< f(t)\,,\, h(t) > = \int f(t)\, h(t)\, dt\,, \tag{B10}$$

and

$$< \theta(\mathbf{x},t)\,,\, \psi(\mathbf{x},t) > = \iint_W \theta(\mathbf{x},t)\, \psi(\mathbf{x},t)\, d\mathbf{x}\, dt\,, \tag{B11}$$

respectively.

Finally, using B5, B7, and B8, equation B6 can be compactly written in the form of a normal equation

$$L^\dagger \, L \, \mathbf{a}(t) = L^\dagger s(\mathbf{x}, T-t) \, . \tag{B12}$$

Even though the source function $s(\mathbf{x}, t)$ in the right hand side of equation B12 is not known, $L^\dagger s(\mathbf{x}, T-t)$ is known as it is approximately equal to the data recorded in the field, i.e.,

$$L^\dagger s(\mathbf{x}, T-t) \approx \mathbf{d}(T-t), \tag{B13}$$

where B13 is not an exact identity because in reality the recorded data is contaminated with noise. Now, if we approximate the first gradient $L^\dagger s(\mathbf{x}, T-t)$ with the recorded data, then the normal equation B12 can be iteratively solved using conjugate gradient. Our numerical experiments (not shown in this report) show that using data contaminated with random or coherent noise as an estimate of the first gradient in the conjugate gradient scheme is possible and conjugate gradient can yield a solution to B12 after 20 to 30 iterations. After solving B12 for $\mathbf{a}(t)$, to get the source image, we must inject and propagate the optimally computed signals $\mathbf{a}(t)$ and then scan the resulting wavefield for the source image.